

Implication of Unobservable State-and-topology Cyber-physical Attacks

Jiazi Zhang, *Student Member, IEEE*, Lalitha Sankar, *Senior Member, IEEE*

Abstract—This paper studies the physical consequences of a class of unobservable state-and-topology cyber-physical attacks in which both state and topology data for a sub-network of the network are changed by an attacker to mask a physical attack. The problem is formulated as a two-stage optimization problem which aims to cause overload in a line of the network with limited attack resources. It is shown that unobservable state-and-topology cyber-physical attacks as studied in this paper can make the system operation more vulnerable to line outages and failures.

Index Terms—Cyber-physical system, false data injection attack, topology, state estimation, two-stage optimization.

I. INTRODUCTION

THE electric power system is a complex cyber-physical system and is monitored by an intelligent which includes: (i) a supervisory control and data acquisition (SCADA) system; and (ii) an energy management system (EMS) that process the SCADA data. Network topology is important system data used in various data processing modules in the EMS. Changes in topology can result from either system incidents or malicious physical attacks; but, in general, such topology alterations can be detected in the cyber layer. However, a sophisticated attacker can launch cyber attacks that alter the topology information in an unobservable manner; furthermore, they can also mask a physical attack via a cyber attack to create a more coordinated attack. Such cyber-physical attacks can result in wrong EMS solutions with potential serious consequences. Therefore, it is instructive to fully understand such attack consequences as a first step to thwart them.

There has been much recent interest in understanding both the physical and cyber security challenges facing the electric power system. While there has been focus on the consequences of physical attacks on the system operation (e.g., [1]), those of cyber as well as coordinated cyber-physical attacks are less understood. In this paper, we introduce a class of unobservable state-and-topology cyber-physical attacks in AC state estimation (SE) and focus on fully understanding its consequences.

A. State of Art

False data injection (FDI) attacks: In [2], Liu *et al.* first introduce a class of FDI attacks on DC SE. In [3], Hug and Giampapa focus on FDI attacks on AC SE and introduce a class of unobservable attacks that are limited to a sub-graph of the networks. They demonstrate that though AC

SE is vulnerable to unobservable FDI attacks, it requires the knowledge of both system topology and states to launch such attacks. More recently, the attacks in [4] by Liang *et al.* study attack consequences by introducing a class of unobservable FDI attacks for AC SE and demonstrate that such attacks can lead to a physical generation re-dispatch and line overflow.

Topology attacks: Unobservable cyber attacks on topology can be of two types: *line-maintaining* and *line-removing*. For a line-maintaining attack, the attacker changes measurements and line status information to make it appear that line that is not in the system is now shown as active at the control center via SCADA data; the opposite is achieved by a line-removing attack. For both line-removing and line-maintaining attacks, an attack can either change only topology data (i.e., state-preserving topology attack) or both state and topology data (i.e., state-and-topology attack). The class of unobservable cyber topology attacks is first introduced in [5]; however, the analysis in [5] is restricted to a subclass of state-preserving line-removing attacks in which an attacker changes topology information of the system without changing the states.

Line-maintaining attacks: This sub-class of topology attacks require both physical line outage and cyber attack to mask the physical topology alteration and have not been studied yet in the literature. In this work, we study the line-maintaining cyber-physical attacks in which both physical and cyber topology are changed by the attacker. In [6], we consider unobservable state-preserving line-maintaining attacks (i.e., only topology data is changed) for which we develop an algorithm using breadth-first search (BFS) to find the smallest sub-network required to launch such an attack. However, changing only topology and not changing states limits the number of feasible lines amenable to attacks and also requires large load shifts at the end buses of a target line. Therefore, in this work, we determine attacks that change both state and topology.

Attack consequences: There has been much focus on effect of attacks on operation costs [7], [8] and electricity markets [9], [10]; in contrast, as in [4], [11], this paper highlights physical system consequences of cyber-physical attacks. For cyber attacks whose goal is to effect electricity market and physical consequences, the attacks can be modeled as two-stage optimization problems where the first stage models the attack design with constraints that capture attacker's limitation and the second stages models the system response (see [7], [8], [11]). In this paper, we also use a two-stage optimization problem to find unobservable state-and-topology cyber-physical attack that can maximize power flow on a chosen line. Furthermore, due to the combination of physical and cyber

J. Zhang and L. Sankar are with the Department of Electrical and Computer Engineering, Arizona State University, Tempe, AZ, 85281 USA e-mail: jzhan188@asu.edu, lalithasankar@asu.edu

attacks, we employ such a two-stage attack twice as detailed in the sequel.

B. Contributions

The contributions of this paper are two-fold. First, we introduce a class of unobservable state-and-topology cyber-physical attacks in which an attacker can change both cyber state and topology data to enable a coordinated physical and cyber attack on AC SE. Such an attack consists of a physical attack to first trip a transmission line, followed by a cyber attack that masks the physical attack. The goal is to overload a chosen line (different from tripped line) while avoiding being detected by both SE and the subsequent modules.

Our attack model also captures the realistic limitation that the attacker can only access a sub-network of the entire power system, and therefore, can take down a line and modify the measurements only inside the sub-network. To this end, we can solve a two-stage optimization problem to determine the attack. However, since both physical attack and the re-dispatch resulting from cyber attack can lead to state changes, two attack vectors are required to enable the above two state changes and ensure the unobservability of the attack. Therefore, we formulate a two-step strategy to determine the attack.

The second contribution of our work is to demonstrate the consequences of the worst cyber-physical attacks determined by the proposed attack strategy on AC SE and AC OPF. We show that the cyber-physical attacks introduced here can successfully lead to line overflows in the IEEE 24-bus system with limited size of attack sub-network and load shifts.

The remainder of this paper is organized as follows. Sec. II introduces the general system model. Sec. III introduces the attack model. Sec. IV presents a two-step attack strategy to identify the worst-case overflow attack. Sec. V analyzes the numerical results for a test system. Sec. VI draws the conclusion of this paper and presents the future works.

II. SYSTEM MODEL

In this section, we introduce the mathematical formulation for the various computational units of power system operation, including system network and topology, state estimation, and optimal power flow. Throughout, we assume there are n_b buses, n_{br} branches, n_g generators, and n_z measurements in the system. In Fig. 1, we illustrate a typical temporal sequence of data processing units in the cyber layer.

A. System Network and Topology

The electric power system can be represented by a graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ where \mathcal{N} and \mathcal{E} are the sets of buses and lines, respectively.

At the control center, SCADA collects line status data as a $n_{br} \times 1$ vector s with entries $s_k \in \{0, 1\}$ for $k \in \{1, \dots, n_{br}\}$ that indicate the on and off status of circuit-breakers on each line. The data is then passed to a topology processor to map the real-time power system topology. Each s corresponds to a specific system topology \mathcal{G} .

B. State Estimation

Consider an $n_z \times 1$ vector z of nonlinear measurements (for AC SE) given as

$$z = h(x, \mathcal{G}) + e \quad (1)$$

where $x = [\theta, V]^T$ is the system state vector, and e is an $n_z \times 1$ noise vector which is independent of x and is modeled as Gaussian distributed with 0 mean and σ_e^2 covariance such that the measurement error covariance matrix is given by $R = \text{diag}(\{\sigma_e^2\}_{i=1}^M)$. The function $h(x, \mathcal{G})$ is a vector of nonlinear functions that describes the relationship between the system states and measurements for a topology \mathcal{G} .

We use weighted least-squares (WLS) AC SE to calculate the θ and V [12]. Subsequent to SE, bad data detector use χ^2 -test to detect bad data and bad data identification use largest normalized residual method to filter the bad data.

C. Optimal Power Flow

The OPF problem can be written as

$$\min C(x) \quad (2)$$

$$\text{s.t. } F(x) = 0 \quad (3)$$

$$T(x) \leq 0 \quad (4)$$

$$x_{\min} \leq x \leq x_{\max} \quad (5)$$

where $x = [\theta, V, P_G, Q_G]^T$ is the optimization vector with voltage angle θ , voltage magnitude V that are both $n_b \times 1$ vectors, and active power generation P_G , reactive power generation Q_G that are both $n_g \times 1$ vectors; $C(\cdot)$ denotes the cost function of x ; $F(\cdot)$ denotes the equivalent constraints (power balance constraints); $T(\cdot)$ denotes the inequivalent constraints (power flow limits).

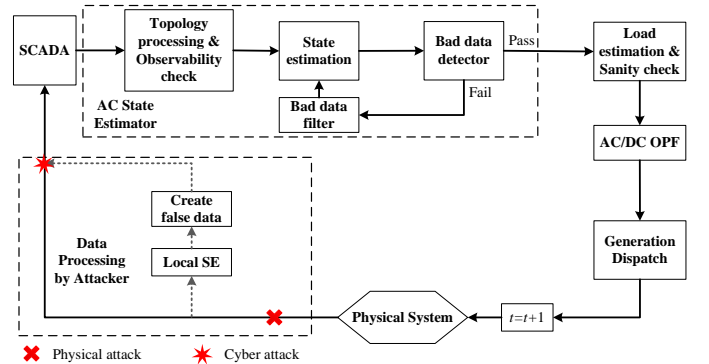


Fig. 1. Temporal Sequence of Data Processing Units in The Cyber Layer within Attack.

III. ATTACK MODEL

The unobservable state-and-topology cyber-physical attack considered here models both a physical attack and a coordinated cyber attack.

We assume the attacker has the following capabilities:

- 1) Attacker has knowledge of the topology $\bar{\mathcal{G}}_0$ of entire network prior to physical attacks.
- 2) Attacker has the capability to launch physical attack, and observe and change measurements only for a sub-graph

\mathcal{S} of \mathcal{G}_0 . The choice of \mathcal{S} is described in detail in the sequel.

- 3) Attacker has the capability to perform SE and compute modified measurements for \mathcal{S} .
- 4) Attacker has knowledge of the capacity and operation cost of every generator in the network.
- 5) Attacker has historic data of load patterns and generation dispatch of the entire network.

We assume that the power system is observable before and after the physical attack.

In this paper, we focus only on physical attacks that target transmission lines. We denote the line that is physically tripped by the attacker as the *switching attack line* and the two end buses of this line as the *switching attack buses*. Assume the switching attack line is line t and the topology prior to the physical attack is \mathcal{G}_0 . The physical line status for line t changes from $s_t = 1$ to $s_t = 0$ after the physical attack and the corresponding physical topology changes to $\bar{\mathcal{G}}$.

In general, a physical attack will be subsequently detected by the topology processing unit in the EMS and the system topology will be updated shortly after the detection. However, a sophisticated attacker can hide such physical attacks by launching an unobservable cyber attack. In the resulting *unobservable cyber topology attack*, the attacker modifies line status as well as related bus measurements to alter the system topology \mathcal{G} to a different “target” topology $\bar{\mathcal{G}} = \{\mathcal{N}, \bar{\mathcal{E}}\}$. Since the attacker’s aim is to hide the topology alteration caused by the physical attack, $\bar{\mathcal{G}}$ should be chosen as \mathcal{G}_0 .

To launch a state-and-topology attack, the attacker injects $n_{br} \times 1$ line status attack vector b and $n_z \times 1$ measurement attack vector a . The attack vector b for line status overrides the physical change on line t ’s status by setting for $b_k = 0$ for $k \neq t$ and $b_k = 1$ for $k = t$. These changes lead to a new system state \bar{x} for the system under attack. This attack modifies (s, z) for topology \mathcal{G} to (\bar{s}, \bar{z}) for topology $\bar{\mathcal{G}}$ such that

$$\bar{s} = s + b, \text{ and } \bar{z} = z + a. \quad (6)$$

In the absence of noise, the measurement attack vector satisfies

$$a = h(\bar{x}, \bar{\mathcal{G}}) - h(x, \mathcal{G}). \quad (7)$$

For nonlinear measurement model and AC SE, we model a sophisticated attacker who attacks measurements and line status data for a sub-graph \mathcal{S} of the network by first estimating the system states \hat{x} inside \mathcal{S} using AC SE. The attacker then chooses a small set of buses in \mathcal{S} to change states from the estimate \hat{x} to $\bar{x} = \hat{x} + c$ such that the measurement vector \bar{z} after cyber attack has entries

$$\bar{z}_i = \begin{cases} z_i, & i \notin \mathcal{I}_{\mathcal{S}} \\ h_i(\hat{x} + c, \bar{\mathcal{G}}), & i \in \mathcal{I}_{\mathcal{S}} \end{cases}. \quad (8)$$

where $\mathcal{I}_{\mathcal{S}}$ denotes the set of measurements inside \mathcal{S} .

We use the following method to identify the sub-graph \mathcal{S} for an unobservable state-and-topology attack. Throughout, we distinguish two types of buses: *load buses* with presence of load and *non-load buses* with no load.

- 1) Use the optimization problem (the details are in the sequel) to determine the load buses from the attack

vector c whose states need to be changed (defined as *center bus*) to enable the attack.

- 2) Include all center buses in \mathcal{S} .
- 3) Extend \mathcal{S} by including all buses and branches connected to the buses inside \mathcal{S} .
- 4) If there are non-load buses on the boundary of \mathcal{S} , extend \mathcal{S} by including all adjacent buses of the non-load boundary buses and the corresponding branches.
- 5) Repeat 4) until all boundary buses of \mathcal{S} are load buses.
- 6) Check if there is a path (actual bus and branch connection) in \mathcal{S} that can connect the two switching attack buses. If such path exists, then \mathcal{S} is the attack sub-graph. If there is no such path, go to Step 7).
- 7) Use BFS method to find the shortest path connecting the two switching attack buses. Include the shortest path in \mathcal{S} . Then this \mathcal{S} is the attack sub-graph.

Steps 1)–5) ensure the boundary buses of \mathcal{S} are load buses with states unchanged. For a non-load bus in \mathcal{S} , since the injection of non-load buses are known to the control center, the attacker should ensure that under an attack, the net injection is equal to the net flow into the bus. Thus, the state changes for non-load buses are dependent on those for the neighboring load buses. Furthermore, the state of a boundary bus j is computed using both measurements inside and outside \mathcal{S} . From (8), if a measurement i for $i \notin \mathcal{I}_{\mathcal{S}}$ is dependent on the j^{th} state, then the corresponding j^{th} entry of the attack vector should satisfy $c_j = 0$ to ensure the attack to be unobservable. Thus, a boundary bus cannot have a state change, and therefore, cannot be a non-load buses.

Steps 6) and 7) ensure that the states of switching attack buses can be estimated with measurements inside \mathcal{S} . To maintain the switching attack line as active in the cyber layer, the attacker needs to modify the line status as well as power flow measurements on the switching attack line and power injection measurements on the switching attack buses. This in turn, requires the attacker to estimate the states of switching attack buses to create the false measurements. However, since this line is physically disconnected, the attacker needs to use an algorithm such as BFS to determine an alternate shortest path connecting the 2 switching attack buses, and thereby estimate the states and changed measurements. In general, state change is required for at least one of the switching attack buses. This bus, thereby, will be included in \mathcal{S} . However, \mathcal{S} may not include the entire physical path. Thus, the attacker needs steps 6) and 7) to complete the path.

IV. ATTACK STRATEGY

In this section, we study the worst-case cyber-physical attacks. We assume the attacks can: (a) physically trip a switching attack line and mask the physical attack with a cyber attack; (b) maximize power flow on a *target line*; and (c) avoid detectability by limiting load shift via changes in measurements. The attack resources available to the attacker may also be limited. We model this limitation by constraining the size of sub-network the attacker has access to. This leads to a constrained optimization problem. As noted before, two attack vectors are needed for the physical and cyber parts of the attack and each optimization problem is described below.

Our two-step optimization problem captures the temporal nature of attack sequence involving a physical attack followed by several cyber attacks.

In Fig. 2, we illustrate this temporal sequence of attack and system events. The system events are periodic and are denoted by S_t for the t^{th} event. At the start of each S_t , data is collected from SCADA and by the end of S_t , *i.e.*, the start of S_{t+1} , data is processed in the EMS. There are 2 attacks instance, A_0 and A_1 to denote the physical and cyber attack events, respectively. We assume the physical attack event A_0 is launched immediately after the start of the 0th system event, *i.e.*, S_0 , and the coordinated cyber attack event A_1 is launched shortly after, but before the start of next system event S_1 . Following this cyber-physical attack pair (A_0, A_1) , the cyber attack is sustained between every two system events to maintain the worst generation dispatch, and thereby, sustain the maximal power flow on the target line. In TABLE I, we denote how the cyber (measured) and physical (actual) data including generation dispatch, system state, topology, and loads vary at all system and attack events.

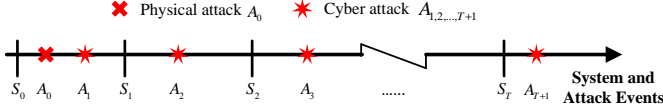


Fig. 2. Time sequence of attack and system events.

TABLE I
PHYSICAL AND CYBER DATA FOR ATTACK AND SYSTEM EVENTS.

System and Attack Event	E_0	A_0	A_1	S_1	A_2	S_2	A_3	...	S_T	A_{T+1}
Generation Dispatch	P_G^0	P_G^0	P_G^0	P_G^*	P_G^*	P_G^*	P_G^*	...	P_G^*	P_G^*
Physical Topology	\bar{G}	G	G	G	G	G	G	...	G	G
Cyber Topology	\bar{G}	$G \parallel \infty \bar{G}$	\bar{G}	\bar{G}	\bar{G}	\bar{G}	\bar{G}	...	\bar{G}	\bar{G}
Physical State	θ_0	θ_0	θ_0	θ^*	θ^*	θ^*	θ^*	...	θ^*	θ^*
Cyber State	θ_0	θ_0	$\theta_0 + c^0$	$\theta^* + c$	$\theta^* + c$	$\theta^* + c$	$\theta^* + c$...	$\theta^* + c$	$\theta^* + c$
Physical Load	P_D	P_D	P_D	P_D	P_D	P_D	P_D	...	P_D	P_D
Cyber Load	P_D	P_D	\bar{P}_D	\bar{P}_D	\bar{P}_D	\bar{P}_D	\bar{P}_D	...	\bar{P}_D	\bar{P}_D

Assume the system topology and the generation at S_0 are \bar{G} and P_G^0 , respectively. From TABLE I, we can see that the system physical topology changes to G after the physical attack. The physical operation states, thereby, change to θ_0 . The attacker then injects cyber attack vector c^0 to change the load pattern from the physical load P_D to the false cyber load \bar{P}_D to mask the physical topology alteration. The physical and cyber loads at attack event A_1 satisfy the following relationships, respectively:

$$P_D = A_{GN} P_G^0 - H_1 \theta_0, \text{ and } \bar{P}_D = A_{GN} P_G^0 - \bar{H}_1 (\theta_0 + c^0) \quad (9)$$

where A_{GN} is $n_b \times n_g$ generator-to-bus connectivity matrix; H_1 and \bar{H}_1 are $n_b \times n_b$ dependency matrices between power

injection and voltage angle for G and \bar{G} , respectively. When subtracting the two equations in (9), the cyber loads are related to the physical loads as

$$\bar{P}_D = P_D + H_1 \theta_0 - \bar{H}_1 (\theta_0 + c^0) \quad (10)$$

The false cyber load \bar{P}_D and topology \bar{G} leads to a system re-dispatch to the optimal generation dispatch P_G^* at S_1 . Since the attacker optimization problem at each step models the system response, such an optimal dispatch will cause maximal power flow on the target line. Following this first cyber attack A_1 , since the generation dispatch changes at S_1 , the physical system states also change to θ^* . To sustain both the optimal dispatch P_G^* and the false cyber topology \bar{G} at the next system event, *i.e.*, S_2 , the attacker needs to maintain the false cyber load \bar{P}_D by injecting another attack vector c at A_2 . Thus, the nodal power balance at attack event A_2 in the cyber layer is:

$$A_{GN} P_G^* - \bar{H}_1 (\theta^* + c) = P_D + H_1 \theta_0 - \bar{H}_1 (\theta_0 + c^0). \quad (11)$$

where the right hand side terms represent the cyber load modified at A_1 . In the following attack events, *i.e.*, A_t , $t = 3, \dots, T$, the attacker can keep injecting c to maintain the false cyber load \bar{P}_D . This in turn ensures that the optimal dispatch and the false cyber topology are maintained at P_G^* and \bar{G} , respectively, and the maximal power flow on the target line is sustained.

To model the cyber-physical attack events A_0 , A_1 , and A_2 between S_0 and S_1 , the optimization problem should capture the power balance relationship shown in (11). However, since the switching attack line is determined by the optimization problem, both H_1 and θ_0 are unknown before solving the problem. On the other hand, for the pure cyber attack events A_2 and A_3 , the power balance in the cyber layer is

$$A_{GN} P_G^* - \bar{H}_1 (\theta^* + c) = P_D + H_1 \theta^* - \bar{H}_1 (\theta^* + c). \quad (12)$$

This is equivalent to the physical power balance as

$$A_{GN} P_G^* - H_1 \theta^* = P_D. \quad (13)$$

Therefore, instead of directly modeling the cyber-physical attack events A_0, A_1 , and A_2 between S_0 and S_1 , we can model the pure cyber attack events A_2 and A_3 between S_1 and S_2 to determine the attack vector c in the first step. Such a c should be subject to bounds on both the attacker's sub-graph size and the load shifts. However, since the new topology H_1 is still not known prior to the optimization, we replace H_1 using the following equations:

$$H_1 \theta^* = A_{KN} P_K \quad (14)$$

$$P_K = \text{diag}(s) \cdot \bar{H}_2 \theta^* \quad (15)$$

where A_{KN} is the $n_b \times n_{br}$ branch-to-bus connectivity matrix, \bar{H}_2 is the $n_{br} \times n_b$ dependency matrix between power flow and voltage angle for \bar{G} , s is the line status vector, $\text{diag}(s)$ represents the diagonal matrix of s , P_K is the $n_{br} \times 1$ power flow vector. In (14), the sum of physical power flows on the set of branches connected to a bus is utilized to calculate the physical power injection at the bus. In (15), the physical power flow vector is represented by the diagonal matrix of line status vector s multiply the cyber power flow vector, *i.e.*, $\bar{H}_2 \theta^*$. That is, if a line t is selected as the switching attack

line, the power flow P_{Kt} on line t is forced to be 0, otherwise, $P_{Kt} = \bar{H}_2(t, :) \cdot \theta$, where $\bar{H}_2(t, :)$ represents the t^{th} row of \bar{H}_2 . With these modifications, we can then model the system and cyber attack events from A_2 through S_2 to A_3 with a two-stage optimization problem, and hence, the switching attack line can also be determined as the solution of the optimization problem. After the switching attack line and the cyber attack vector c are both determined, the attack sub-graph \mathcal{S} can be identified with the process stated in Section III. The details of this problem is described in Subsection IV-A.

In the second step, we focus on the attack vector c^0 at A_1 . We again use a two-stage optimization problem to determine the c^0 such that the optimal generation dispatch for this problem is forced to be same as that in Step 1. We, henceforth, define the attack vector solved in the second step as the *initial attack vector*. The details of the second step is introduced in Subsection IV-B.

The attack vectors c and c^0 are both DC attack vectors that can be detected by AC SE. Thus, to ensure the unobservability of the attacks, the attacker should construct two AC attacks with c and c^0 . This procedure is introduced in Subsection IV-C.

A. Step 1: Maximize Power Flow on A Line

In Step 1, we introduce a two-stage optimization problem to determine the attack vector c and the switching attack lines such that the target line l in the attacker's sub-graph \mathcal{S} has maximal power flow subject to specific constraints as explained in the sequel. The two-stage optimization is given as

$$\max P_{Kl} - \zeta \|c_{\mathcal{L}}\|_0 \quad (16)$$

$$\text{s.t.} \quad \sum_{k=1}^{n_{br}} (1 - s_k) = N_T, \quad s_k \in \{0, 1\} \quad (17)$$

$$\|c_{\mathcal{L}}\|_0 \leq N_0 \quad (18)$$

$$-\tau P_D \leq \bar{H}_1 (\theta^* + c) - A_{KN} P_K^* \leq \tau P_D \quad (19)$$

$$\{\theta^*, P_G^*, P_K^*\} = \arg \left\{ \min_{\theta, P_G, P_K} \sum_{g=1}^{n_g} C_g(P_{Gg}) \right\} \quad (20)$$

$$\text{s.t.} \quad A_{GN} P_G - A_{KN} P_K = P_D \quad (\lambda) \quad (21)$$

$$P_K = \text{diag}(s) \cdot \bar{H}_2 \theta \quad (22)$$

$$-P_K^{\max} \leq \bar{H}_2 (\theta + c) \leq P_K^{\max} \quad (\mu^-, \mu^+) \quad (23)$$

$$P_G^{\min} \leq P_G \leq P_G^{\max} \quad (\alpha^-, \alpha^+) \quad (24)$$

where $C_g(\cdot)$ is the cost function for generator g ; P_G is $n_g \times 1$ active power generation vector with maximum and minimum limit P_G^{\max} and P_G^{\min} , respectively; P_K is $n_{br} \times 1$ physical power flow vector with thermal limit P_K^{\max} ; λ is $n_b \times 1$ dual variable vector of constraint (21); μ^\mp are $n_{br} \times 1$ dual variable vectors of constraint (23), respectively; α^\mp are $n_b \times 1$ dual variable vectors of constraint (24), respectively; \bar{H}_1 is $n_b \times n_b$ dependency matrix between power injection and voltage angle for \mathcal{G} ; \bar{H}_2 is $n_{br} \times n_b$ dependency matrix between power flow and voltage angle for \mathcal{G} ; P_D is $n_b \times 1$ physical load vector, which has maximum load shift percentage τ ; ζ is the weight of the norm of attack vector c ; \mathcal{L} represents the set of load buses; N_0 is the maximum number of load-buses that can

be attacked; N_T is the maximum number of switching attack lines.

The goal of the attack in (16) is a multi-objective problem which includes maximizing the power flow on the target line l to create an overflow, while minimizing the l_0 -norm of the attack vector, *i.e.*, minimizing the attack sub-graph size. The power flow on l is maximized along the direction of the power flow prior to attack. In the first stage, constraints (17)–(19) model the following attacker limitations: (i) only up to N_T switching attack lines can be physically tripped; (ii) alter up to N_0 load-bus states; and (iii) limit cyber load shifts to at most τP_D ; respectively. The second stage optimization represents DC OPF, whose aim is to minimize operation cost in (20), subject to power balance constraints in (21) and (22), thermal limit constraint in (23), and generation limit constraint in (24).

This two-stage optimization problem is nonlinear and non-convex. For tractability, we modify several constraints.

Constraint (22) is a nonlinear constraint which includes the product of binary variable s and continuous variable θ . It can be replaced by a linear form as follows

$$\begin{cases} -P_K + \bar{H}_2 \theta^* \leq M_1 (1 - s) & (\beta^-) \\ P_K - \bar{H}_2 \theta^* \leq M_1 (1 - s) & (\beta^+) \\ -P_K \leq M_1 \cdot s & (\gamma^-) \\ P_K \leq M_1 \cdot s & (\gamma^+) \end{cases} \quad (25)$$

where β^\pm and γ^\pm are $n_{br} \times 1$ dual variable vectors for the corresponding constraints and M_1 is a large number.

Constraint (18) is an l_0 -norm constraint on the attack vector, which is nonlinear and non-convex. It can be relaxed to a corresponding l_1 -norm constraint as:

$$\|c_{\mathcal{L}}\|_1 = \sum_{n \in \mathcal{L}} |c_n| \leq N_1. \quad (26)$$

However, constraint (26) is still nonlinear. We, thus, linearize it as follows:

$$c_n \leq u_n, \quad -c_n \leq u_n, \quad \sum_{n \in \mathcal{L}} u_n \leq N_1. \quad (27)$$

where u is $n_{load} \times 1$ non-negative slack variable vector.

Once the attack vector determined by s and c is given in the first stage optimization problem, the second stage DCOPT problem (20)–(24) and (25) is then convex. The second stage optimization problem can then be replaced by its Karush-Kuhn-Tucker (KKT) optimality conditions as follows:

$$\begin{aligned} & \nabla \left(\sum_{g=1}^{n_g} C_g(P_{Gg}^*) \right) + \lambda^T \cdot \nabla (A_{GN} P_G^* - A_{KN} P_K^* - P_D) \\ & + [\mu^-; \mu^+]^T \cdot \nabla \left(\begin{bmatrix} -\bar{H}_2 (\theta^* + c) \\ \bar{H}_2 (\theta^* + c) \end{bmatrix} - \begin{bmatrix} P_K^{\max} \\ P_K^{\max} \end{bmatrix} \right) \\ & + [\alpha^-; \alpha^+]^T \cdot \nabla \left(\begin{bmatrix} -P_G^* \\ P_G^* \end{bmatrix} - \begin{bmatrix} -P_G^{\min} \\ P_G^{\max} \end{bmatrix} \right) \\ & + [\beta^-; \beta^+]^T \cdot \nabla \left(\begin{bmatrix} -P_K^* \\ P_K^* \end{bmatrix} + \begin{bmatrix} \bar{H}_2 \theta^* \\ -\bar{H}_2 \theta^* \end{bmatrix} - M_1 \cdot \begin{bmatrix} 1-s \\ 1-s \end{bmatrix} \right) \\ & + [\gamma^-; \gamma^+]^T \cdot \nabla \left(\begin{bmatrix} -P_K^* \\ P_K^* \end{bmatrix} - M_1 \cdot \begin{bmatrix} s \\ s \end{bmatrix} \right) = 0 \end{aligned} \quad (28)$$

$$\text{diag}([\mu^-; \mu^+]) \cdot \left(\begin{bmatrix} -\bar{H}_2 (\theta^* + c) \\ \bar{H}_2 (\theta^* + c) \end{bmatrix} - \begin{bmatrix} P_K^{\max} \\ P_K^{\max} \end{bmatrix} \right) = 0 \quad (29)$$

$$\text{diag}([\alpha^-; \alpha^+]) \cdot \left(\begin{bmatrix} -P_G^* \\ P_G^* \end{bmatrix} - \begin{bmatrix} -P_G^{\min} \\ P_G^{\max} \end{bmatrix} \right) = 0 \quad (30)$$

$$\text{diag}([\beta^-; \beta^+]) \cdot \left(\begin{bmatrix} -P_K^* \\ P_K^* \end{bmatrix} + \begin{bmatrix} \bar{H}_2 \theta^* \\ -\bar{H}_2 \theta^* \end{bmatrix} - M_1 \cdot \begin{bmatrix} 1-s \\ 1-s \end{bmatrix} \right) = 0 \quad (31)$$

$$\text{diag}([\gamma^-; \gamma^+]) \cdot \left(\begin{bmatrix} -P_K^* \\ P_K^* \end{bmatrix} - M_1 \cdot \begin{bmatrix} s \\ s \end{bmatrix} \right) = 0 \quad (32)$$

$$[\mu^-; \mu^+; \alpha^-; \alpha^+; \beta^-; \beta^+; \gamma^-; \gamma^+] \geq 0 \quad (33)$$

where constraint (28) is the partial gradient optimal condition, (29)–(32) are the complementary slackness constraints, (21)–(24) and (25) are the primal feasibility constraints, and (33) represents the dual feasibility constraints.

Particularly, the complementary slackness constraints (29)–(32) are nonlinear since they include product of continuous variables. We then linearize them by introducing new binary variables δ_{μ^\pm} , δ_{α^\pm} , δ_{β^\pm} , and δ_{γ^\pm} . For instance, constraint (29) can be rewritten as

$$\begin{cases} \mu^- - M \cdot \delta_{\mu^-} \leq 0 \\ \bar{H}_2(\theta^* + c) + P_K^{\max} \leq M(1 - \delta_{\mu^-}) \\ \mu^+ - M \cdot \delta_{\mu^+} \leq 0 \\ -\bar{H}_2(\theta^* + c) + P_K^{\max} \leq M(1 - \delta_{\mu^+}) \end{cases} \quad (34)$$

where M is a large positive number. Constraints (30)–(32) can be linearized using the same method. Particularly, for the linearized forms of constraints (31) and (32), M_1 and M are different values and $M_1 \ll M$.

Using the approximate relaxation for the various constraints as detailed above, we obtain the following equivalent single-stage mixed-integer linear problem with the objective

$$\max P_{Kl} - \zeta \sum_{n \in \mathcal{L}} u_n \quad (35)$$

subject to (17), (19), (21), (23)–(25), (27), (28), (33), (34), and the linearized forms of constraints (30)–(32). Note that since no real-time data is required in the above optimization problem, the attacker can solve this step offline to determine the switching attack line to trip and the attack vector c .

B. Step 2: Determine Initial Attack Vector

In this step, we determine the attack vector at events A_1 . As stated earlier, c^0 , the attack vector at A_1 , is chosen to ensure that the resulting load shifts lead to the optimal dispatch solved in Step 1, i.e., P_G^* . To this end, we use a two-stage optimization problem similar to Step 1 to determine c^0 . Note that since the switching attack line and attack sub-graph are both determined in Step 1, the dependency matrix between power injection and voltage angle, i.e., H_1 for the physical topology \mathcal{G} at A_1 is known to the attacker, and the cyber loads are given by (10).

As stated in Section III, the attacker only has access to the measurements inside \mathcal{S} . Thus, the attacker cannot directly obtain the whole system physical states θ_0 . However, assuming \mathcal{B}_S and \mathcal{B}_{NS} represents the set of buses inside and outside \mathcal{S} , respectively. The vector of cyber loads resulting from an unobservable attack satisfies the following relationship:

$$\bar{P}_D = \begin{bmatrix} P_D^{\mathcal{B}_S} + H_1^{\mathcal{B}_S} \theta_0^{\mathcal{B}_S} - \bar{H}_1^{\mathcal{B}_S} (\hat{\theta}_0^{\mathcal{B}_S} + c^0) \\ P_D^{\mathcal{B}_{NS}} \end{bmatrix} \quad (36)$$

where $P_D^{\mathcal{B}_S}$ is the vector of physical loads for all buses inside \mathcal{S} , $P_D^{\mathcal{B}_{NS}}$ is that for all buses outside \mathcal{S} , $H_1^{\mathcal{B}_S}$ and $\bar{H}_1^{\mathcal{B}_S}$ represents the sub-matrices of H_1 and \bar{H}_1 for the set of buses inside \mathcal{S} , respectively. For the physical system states $\theta_0^{\mathcal{B}_S}$, attacker uses the estimated states $\hat{\theta}_0^{\mathcal{B}_S}$, to compute (36).

The two-stage optimization can be written as follows:

$$\min \|c_{\mathcal{L}}\|_0 \quad (37)$$

$$\text{s.t.} \quad -\tau P_D^{\mathcal{B}_S} \leq \bar{H}_1^{\mathcal{B}_S} (\hat{\theta}_0^{\mathcal{B}_S} + c^0) - H_1^{\mathcal{B}_S} \hat{\theta}_0^{\mathcal{B}_S} \leq \tau P_D^{\mathcal{B}_S} \quad (38)$$

$$\|c_{\mathcal{L}}\|_0 \leq N_0 \quad (39)$$

$$\{\bar{\theta}^*, P_G^*\} = \arg \left\{ \min_{\bar{\theta}, P_G} \sum_{g=1}^{n_g} C_g(P_{Gg}) \right\} \quad (40)$$

$$\begin{aligned} \text{s.t.} \quad & A_{GN} \cdot P_G - \bar{H}_1 \cdot \bar{\theta} \\ & = \begin{bmatrix} P_D^{\mathcal{B}_S} + H_1^{\mathcal{B}_S} \hat{\theta}_0^{\mathcal{B}_S} - \bar{H}_1^{\mathcal{B}_S} (\hat{\theta}_0^{\mathcal{B}_S} + c^0) \\ P_D^{\mathcal{B}_{NS}} \end{bmatrix} \quad (\lambda) \end{aligned} \quad (41)$$

$$-P_K^{\max} \leq \bar{H}_2 \bar{\theta} \leq P_K^{\max} \quad (\mu^-, \mu^+) \quad (42)$$

$$P_G = P_{G1}^* \quad (\alpha) \quad (43)$$

where P_{G1}^* is $n_g \times 1$ optimal generation vector solved in Step 1; α is $n_g \times 1$ dual variable vector of constraints (43). The objective (37) is to minimize the l_0 -norm of the attack vector. Constraint (37) represents load shift limitation. Constraints (40)–(43) represent the second stage DCOPF problem, which guarantees that the attack vector selected in the first stage leads to the optimal dispatch P_{G1}^* . The l_0 -norm constraint can be relaxed to a linearized l_1 -norm constraint as (26). The objective can be represented as $\sum_{n \in \mathcal{L}} u_n$. This problem can then be converted to a single stage optimization problem using methods similar to those as in detailed Step 1.

C. Implementation

The method to construct an unobservable AC attack with a DC attack vector has been introduced in [11] for FDI attacks without topology alteration. In this paper, we focus on constructing AC unobservable cyber-physical attacks. The procedure is as follows:

- 1) Solve the Step 1 optimization offline to obtain the switching attack line t and the attack vector c .
- 2) Identify the attack sub-graph \mathcal{S} with c and line t .
- 3) Launch the physical attack on the switching attack line.
- 4) Perform local SE inside \mathcal{S} with slack bus chosen as one arbitrary load bus in \mathcal{S} to obtain $\hat{\theta}_0^{\mathcal{B}_S}$;
- 5) Solve the Step 2 optimization problem to obtain c^0 ;
- 6) For all load buses m inside \mathcal{S} , set $\bar{\theta}_{0m} = \hat{\theta}_{0m} + c_{0m}^0$;
- 7) For all non-load buses, since the net injections are not changed, the nodal balance equations for each non-load buses are

$$A_{GN}^m P_G - V_m \sum_{j \in \mathcal{B}r_m} V_j (G_{mj} \cos \theta_{mj} + B_{mj} \sin \theta_{mj}) = 0 \quad (44)$$

$$A_{GN}^m Q_G - V_m \sum_{j \in \mathcal{B}r_m} V_j (G_{mj} \cos \theta_{mj} - B_{mj} \sin \theta_{mj}) = 0 \quad (45)$$

where A_{GN}^m represents the m^{th} row of A_{GN} , Q_G represents the reactive power generation vector, $G_{mj} + iB_{mj}$ is the $(m, j)^{\text{th}}$ entry of the bus admittance matrix, and $\theta_{mj} = \theta_m - \theta_j$ is the voltage angle difference between bus m and j , $\bar{B}r_m$ is the set of branches connecting to bus m for \bar{G} . These equations can be solved iteratively with Newton-Raphson method.

- 8) After updating the cyber states for the non-load buses, using equation (8) to calculate the AC attack for A_1 .
- 9) Repeat Steps 4)–7) (without solving Step 2 optimization) to construct AC attacks with c for A_2, \dots, T .

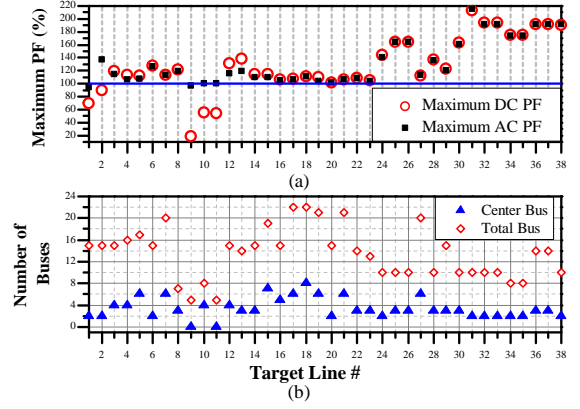
V. NUMERICAL RESULTS

In this section, we test the effect of attacks designed with the two-step attack strategy for a nonlinear system model. The test system is the IEEE 24-bus reliable test system (RTS). We assume: (i) the system is operating under optimal power flow; and (ii) the loads of the system are constant and are equivalent to the historic load data that is assumed to be known to the attacker. To model realistic power systems, we assume that there are congestions prior to the attack and the attacker chooses one congested line as target to maximize power flow. We use MATPOWER to run AC power flow and AC OPF. The optimization problem is solved with CPLEX.

A. Solution for the attack designed with the attack strategy

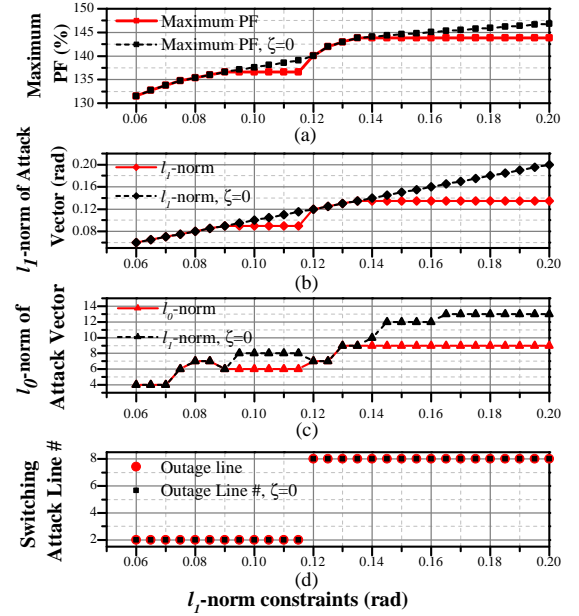
The solution of the unobservable topology attack determined by the two-step attack strategy is tested in this subsection. In order to understand the worst-case effect of attacks, we assume there is a line congested prior to the physical attack. This is achieved in simulation by reducing the line rating to 95% of the base case power flow (apparent power) to create congestion. We exhaustively test all 38 lines as targets in the system and let ζ , the weight for the l_1 -norm term in the objective in (35), be 1% of the original power flow on each target line. Fig. 3 illustrates the maximal power flow (PF) and attack size (# of buses in sub-graph) for load shift bounds $\tau = 10\%$, total lines to physically attack $N_T = 1$, and the l_1 -norm constraint $N_1 = 0.06$. The plot in Fig. 3(a) indicate the flow attack end of system event S_1 using attack vector from event A_1 . In Fig. 3(a), we compare the physical power flow (apparent power, we denote it as AC PF) in each line to the power flow solved in linear model (we denote it as DC PF). In Fig. 3(b), we plot the number of center buses, *i.e.*, l_0 -norm of the attack vector, and the total number of buses inside the attack sub-graph for each target line.

From Fig. 3(a) we can observe that the attack vector determined by the two-stage optimization problem cause overflows in 33 target lines in linear model, *i.e.*, 86.84% of the attacks are successful. For all such successful attacks, using the attack vector to construct an attack in the nonlinear model, in Fig. 3(a), the AC PF in each line tracks DC PF solved with the attack strategy. In particular, 2 cases with target lines 9 and 11, respectively, have no center buses, *i.e.*, for these lines the state-preserving attacks introduced in [6] suffice. In Fig. 3(b), we can observe that 72.73% of the successful attacks can be launched inside a sub-graph with less than 16 total buses.



(a) Maximum DC & AC PF (b) Number of Center Buses & Total Buses inside Attack Sub-graph v.s. Target Line.

Fig. 3. Summary of all 38 target lines under with $\tau = 10\%$ and $N_1 = 0.06$.



(a) Maximum PF (b) l_1 -norm of attack vector (c) l_0 -norm of attack vector (d) the switching attack line v.s. the conditional l_1 -norm constraints. Fig. 4. Target Line 12 (Connecting Bus 8 - Bus 9) with $\tau = 10\%$.

In Fig. 4, we illustrate the effect of the l_1 -norm constraint on the maximal power flow (Fig. 4(a)), the l_1 - and the l_0 -norms of the attack vector (Figs. 4(b) and 4(c), respectively), and the switching attack line (Fig. 4(d)) for target line 12 solved with Step 1 optimization. In each sub-figure, we illustrate the two solutions: one with ζ (red) and the other without ζ (black) in the objective function. From Figs. 4 (a)–(c), we can see that for the solutions without l_1 -norm in objective (*i.e.*, $\zeta = 0$) as the l_1 -norm constraint is relaxed, the maximal target line power flow as well as the l_1 - and l_0 -norms of attack vector also increase. In contrast, for plots with the l_1 -norm in the objective, the l_1 -norm ensures that the vector with the smallest number of center buses is chosen. This in turn implies that when the l_1 -norm in the objective is tight, the resulting power flow may be smaller than that obtained without such a constraint. These differences are illustrated in Figs. 4(a)–(c). In Fig. 4(d), we demonstrate that

the switching attack line chosen by the optimization problem changes from line 2 to line 8 as the l_1 -norm constraint is relaxed. In general, tripping line 8 requires a large load shift, and thus, is only possible for larger l_1 -norm constraint as then the cyber load changes can be distributed over a larger number of load buses in the sub-graph.

B. Consequences of the attack in the nonlinear model

In this subsection, we select a typical case to demonstrate the consequence of the unobservable state-and-topology cyber-physical attack determined by the attack strategy in the nonlinear system model. In this case, the target line is line 12 with $\tau = 10\%$, $N_T = 1$, and $N_1 = 0.06$. Under this condition, the switching attack line is line 2.

For the chosen target line, after launching the physical attack at A_0 and injecting the initial cyber attack constructed with c^0 at A_1 , the active power generation dispatch for generators at bus 7 and 13 change from 215.69 MW and 230.96 MW to 200.69 MW and 245.67 MW, respectively (the dispatch of other generators remain unchanged). In the following events, as the attacker continues to inject the AC attacks constructed with attack vector c (determined by Step 1 optimization), the active power generation for these two set of generators are maintained at these values. Fig. 5 demonstrates the cyber and physical power flow variation during 20 system events. From Fig. 5, we can observe that once the active power generation dispatch changes to the optimal dispatch and remains unchanged in the subsequent system events, the physical overflow in the target line will be maintained by injecting the AC attack constructed with attack vector c . The heat accumulation may eventually cause this line to overheat and then trip offline all the while remaining unobservable to the control center.

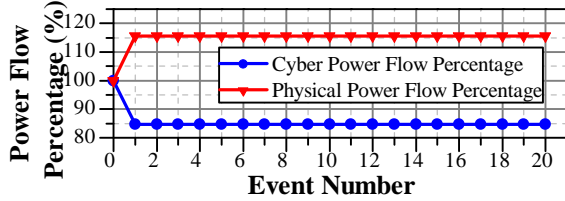


Fig. 5. Power Flow Variation on Line 12 During 20 System Events.

We compare the load shifts caused by the attacks in both the linear and nonlinear system models and find that the load shifts in the nonlinear system model track those in the linear system model for most of the successful attacks. The only two exceptions are for lines 13 (*i.e.*, 20% load shift on a bus) and 23 (*i.e.*, 15% load shift on a bus).

For the successful attacks with other target lines, we observe similar attack consequences in the nonlinear model.

VI. CONCLUDING REMARKS

In this paper, we have introduced a class of unobservable topology attacks in which both topology data and states for a sub-graph of the network are changed by an attacker. We have proposed a two-step attack strategy to maximize the power flow on a target line subject to constraints on limited size of

attack sub-graph and limited load shifts. We have shown that attacks designed with the proposed two-step attack strategy can cause physical line overloads in the IEEE 24-bus RTS even when the attack is subject to bounds on changes in load, for both linear and nonlinear models. The proportion of the successful attacks in the nonlinear system model is 86.84%, which shows the vulnerability of the system to such attacks.

A potential countermeasure is to use historical data to forecast and predict expected generation dispatch. The cyber load patterns created by the attack will in general be different from the normal load shift patterns that lead to the same dispatch plan. Thus, such forecasting can lead to detection of anomalies in both loads and dispatch.

An important extension to study is to understand the impact of attacks when the attacker has access to topology and generation data only for a sub-network. While our attack model restricts data changes to a sub-graph, it still requires the attacker to have knowledge of the complete system topology and generation data. Yet another avenue is to study the worst-case attacks that trip multiple switching attack lines and maximize power flow on multiple target lines.

REFERENCES

- [1] J. Salmeron, K. Wood, and R. Baldick, "Analysis of electric grid security under terrorist threat," *Power Systems, IEEE Transactions on*, vol. 19, no. 2, pp. 905–912, May 2004.
- [2] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM Conference on Computer and Communications Security*, ser. CCS '09, Chicago, Illinois, USA, 2009, pp. 21–32.
- [3] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.
- [4] J. Liang, O. Kosut, and L. Sankar, "Cyber-attacks on ac state estimation: Unobservability and physical consequences," in *IEEE PES General Meeting*, Washington, DC, July 2014.
- [5] J. Kim and L. Tong, "On topology attack of a smart grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, Washington, DC, February 2013, pp. 1–6.
- [6] J. Zhang and L. Sankar, "Implementation of unobservable state-preserving topology attacks," in *North American Power Symposium (NAPS), 2015, Accepted*, Oct 2015, pp. 1–6.
- [7] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *Smart Grid, IEEE Transactions on*, vol. 2, no. 2, pp. 382–390, June 2011.
- [8] —, "Quantitative analysis of load redistribution attacks in power systems," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 9, pp. 1731–1738, Sept 2012.
- [9] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.
- [10] L. Jia, J. Kim, R. J. Thomas, and L. Tong, "Impact of data quality on real-time locational marginal price," *IEEE Trans. Power Systems*, vol. 29, no. 2, pp. 627–636, 2014.
- [11] J. Liang, O. Kosut, and L. Sankar, "Consequences and vulnerability analysis of false data injection attack on power system state estimator," *IEEE Trans. Power Systems*, under review, 2015.
- [12] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York: CRC Press, 2004.